



Show Something: Intelligent Shopping Assistant Supporting Quick Scene Understanding and Immersive Preview

Hao Dou^{1(✉)}, Zhinan Li¹, Minghao Cai¹, Kelvin Cheng²,
Soh Masuko², and Jiro Tanaka¹

¹ Waseda University, Kitakyushu, Japan

{douhao, lizhinan}@fuji.waseda.jp,

mhcai@toki.waseda.jp, jiro@aoni.waseda.jp

² Rakuten Institute of Technology, Rakuten, Inc., Tokyo, Japan

{kelvin.cheng, so.masuko}@rakuten.com

Abstract. In this paper, we introduce an intelligent shopping assistant system supported by quick scene understanding and augmented-reality 3D preview. By understanding the scene that users are looking at, and using the detected scene information, our system recommends related products that are not in the current scene and which could potentially interest users. With the help of existing speech recognition techniques, our system extracts users' voice command and keywords, and provides responses in real-time, which allows users to search and filter specific products just by using voice. After finding the potential target products, our system provides users with an augmented-reality preview experience. It automatically brings products to the suitable space in front of the users by using life-size three-dimensional virtual products and spatial understanding. Users can also use two-hand gestural manipulation to operate the virtual products. Through our system, users can obtain products that are strongly related to the current environment, and intuitively preview the products in the current scene by automatic placement and two-hand manipulation to make shopping decisions.

Keywords: Immersive shopping · Intelligent system · Context awareness

1 Introduction

Online shopping systems have developed rapidly due to its flexibility and its ability to connect to large information networks, enabling users to purchase products whenever and wherever they want [1]. However, current online shopping systems are still limited. The recommendation functions of these shopping systems are usually limited to purchase history or search history. Image-based searches typically focus on methods on finding the same class of products given an image of the scene. The dominant objects from the image are extracted, and identical or similar products are returned from an e-commerce platform. However, users do not necessarily want to purchase the exact same item as in the scene. For example, when a user looks at a sink in the kitchen, the user may want to purchase detergent, cutlery, and other kitchen items instead of buying another sink.

© Springer Nature Switzerland AG 2019

S. Yamamoto and H. Mori (Eds.): HCII 2019, LNCS 11570, pp. 205–218, 2019.

https://doi.org/10.1007/978-3-030-22649-7_17

In addition, current web shopping systems are mostly limited to showing 2D information and usually cannot let users see a digital preview of items in the physical world instantly. Recently, evolving technologies that make use of Augmented-Reality (AR) and computer vision techniques seek to automate tasks that enhance the human visual system can bring new opportunities for intelligent shopping environments.

This work aims to realize an intelligent shopping assistant system supported by quick scene understanding. Combining voice control, automatic placement and two-hand manipulation, this system can also provide users with an immersive augmented-reality preview experience. Using our system, users can get high-quality product recommendations based on the current scene and make quick shopping decisions through products filtering and previewing.

2 Goal and Approach

Our work aims to present an intelligent shopping assistant system which supports quick scene understanding to recommend the user scene-related products which users do not own themselves, and help users make shopping decision quickly. Our system also supports intelligent voice control, an immersive augmented-reality preview experience, as well as providing support for two-hand manipulation with automatic placement.

The quick scene understanding is aimed to recommend products intelligently according to the current scene of the users. Users only need to look at the current scene, and the system will capture and analyze the current scene rapidly and provide result precisely. This can be realized by using scene recognition method. With this method, the picture received from the camera can be analyzed, and objects in the picture are recognized, and finally the possible scene is inferred.

Voice control is used to integrate the intelligent shopping experience for the users. It allows users to say exactly what they want by using voice commands, allowing the use of multiple keywords in one sentence, and filtering out products by using adjectives. We realized this by using text analysis [2] method. It can extract more than one keyword from a sentence accurately. Adjectives can also be extracted to search more precisely, such as “blue shirt” or “big table”. It should also allow the users to filter the products, such as “I want to buy a pair of shoes.”, “Filter, I prefer Nike.” The system will search for all the shoes first and then search for Nike shoes.

Two-hand manipulation will allow users to interact with the preview of the products in augmented-reality. The search results should contain 3D virtual objects that can be dragged into the real world. Users can use two-hand manipulation [3] to move and rotate the virtual objects.

The system should automatically place the virtual objects according to the space availability in the physical world. After scanning the scene, the system will find enough space to place the object and place them automatically. If there is not enough space, the products which do not fit in the space well will not be placed.

3 Intelligent Shopping Assistant

3.1 Scene Understanding

Current systems that search through images typically focus on methods that search for products that are identical to products within the image. They extract specific objects from the image and try to search the exact same item in the e-commerce platforms. However, sometimes users do not necessarily want to buy the same item as in the scene. In an image, the ambiance of the users' environment can also be treated as a source of rich contextual information [4]. When people visit various places throughout the day, many of the impulses that drive shopping are generated by objects in the environment. For example, when users see a sink in the kitchen, they may want to buy some detergent, cutlery, and other kitchen items instead of buying another sink.

In this system, we introduce a way to search for products within the scene. By understanding the scene that users are currently looking at, and using the detected scene information, our system recommends related products which could potentially interest users (see Fig. 1).



Fig. 1. The system recommends relevant products to the user according to the current scene in the user's environment. Scene recognition results: a kitchen with a sink and a mirror.

A scene is a view of a real-world environment where users are physically located, and which contains multiples surfaces and objects, organized in a meaningful way. With the help of computer vision techniques [5], the scene can be understood quickly, after generating an overall description of the scene (e.g. kitchen, or restroom), dominate objects that the system is able to recognize (e.g. desktop computer, or desk) from the ambient environment is included. We use the keyword of the current scene

(e.g. kitchen) to search the products in the e-commerce platform that have the related categories as the dominate objects in the scene (instead of the dominate objects themselves). In this way, we recommend users with products that are strongly associated with the current scene.

For example, a user scans the current environment through our system and expects to receive product recommendations. Our system scans to find out that the current environment is a kitchen, and the current environment has a dominate object, a microwave oven. The system first obtains the search keyword: “kitchen”. Next, the e-commerce platform searches for the category relationship of the dominant objects. For example, the category of the microwave oven is: “microwave oven”, it has a parent category “appliances.” We will use the keyword “kitchen” to search for the category “appliances” and exclude the category “microwave oven” from the search results. Users will get search results for kitchen appliances such as coffee machines, bread machines, and electronic ovens. If there are multiple dominant objects, we will use each dominant object and repeat the above search process. The search results will be combined and return to the users. In this way, we recommend users products that are closely related to the current scenario that users may not already own.

Figure 2 shows the process of searching products by scene: (a) The user issues the voice command “show something”, the system then takes a photo, and the photo is sent to the scene recognition server. (b) The scene recognition server analyzes the keyword “kitchen” in the current scene and the dominant object “microwave oven” in the scene and sends the dominant object “microwave oven” to our e-commerce category API. (c) The directory API searches for the category and parent category of the “microwave oven” (d) The system searches for the current scene keyword “kitchen” in the parent category “appliances” and exclude the category “microwave oven” and returns the recommended products to the user.

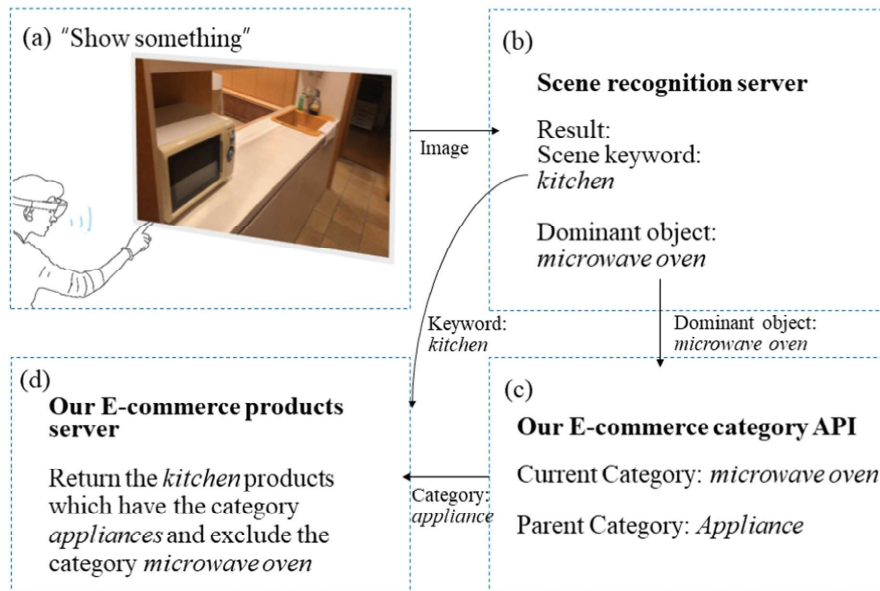


Fig. 2. The process of searching products by scene:

These recommended items are displayed in augmented-reality in a virtual panel that floats in the real world. Users can visually see the links between recommended products and the real world, and view the details of the recommended items, or directly filter the recommended items by voice command and quickly find the products they are interested in.

3.2 Voice Command

To be able to get specific products, a method of interaction is required. In this section, we introduce an intelligent voice interaction approach.

Previous intelligent voice control is quite simple. Many intelligent speech recognition systems can only recognize simple sentences and cannot understand the meaning of the user's voice instructions through context [6]. Users can only ask the system to search specific products instead of searching detail information. If users are not satisfied with the search result, users can not add more filter conditions to filter the searching result. Instead of using this kind of voice control, users prefer using a keyboard to type the searching conditions through the keyboard.

To improve on the existing voice control methods, we propose an intelligent voice control approach, which allows users to use voice command to search products with different attributes and filter the products by voice command.

Table 1. Nouns extracted from example instructions

Instructions	Keyword
<i>"I want to buy a book."</i>	<i>book</i>
<i>"I would like a bag."</i>	<i>bag</i>
<i>"A table is what I need."</i>	<i>table</i>
<i>"Would you please show me a chair?"</i>	<i>chair</i>
<i>"Show a lamp."</i>	<i>lamp</i>

In order to search products through voice command, we need to analyze users' voice instruction and extract the core content of the instruction. We used text analysis method to get the users' keywords from the sentence. This method first analyzes the user's voice instruction and classify the words by their part-of-speech. We extract the nouns as the keywords of the instruction (Table 1).

To search products in a more precise way, we allow users to search with multiple nouns. In this case, more than one noun can be extracted. The nouns can be used as the qualifiers or other keywords (Table 2).

Table 2. Multiple nouns extracted from example instructions

Instructions	Keyword
<i>"I want to buy a children book."</i>	<i>children book</i>
<i>"Show a lady backpack."</i>	<i>lady backpack</i>
<i>"I would like a floor lamp."</i>	<i>floor lamp</i>
<i>"I want a notebook and a pencil."</i>	<i>notebook, pencil</i>

We also made the adjectives qualifiers, so that the users can search the products with different attributes (Table 3).

Table 3. Nouns (Modified by adjectives) extracted from example instructions

Instructions	Keyword
"I want to buy a blue chair."	blue chair
"I want a big table."	big table
"Show the latest laptop."	latest laptop
"I'd like a lotion for dry skin."	lotion, dry skin
"I am sensitive skin and I want to buy a sunscreen."	sensitive skin, sunscreen

In addition, users can say multiple sentences to filter the products. Users can first say a sentence to search a product, they can then add more filter conditions to narrow down the search results. The more filter conditions users add, the more accurate the results are (Table 4). Users can also say "Go back" to recall their filter operation.

Table 4. Nouns (Filtered by conditions) extracted from example instructions

Instructions	Keyword
"I want to buy a pair of sneakers.", "Filter, I love Adidas."	Adidas sneakers
"I want a table.", "Filter, I want cheaper ones."	cheaper table

With the above search methods, users are able to perform accurate products searching (see Fig. 3).

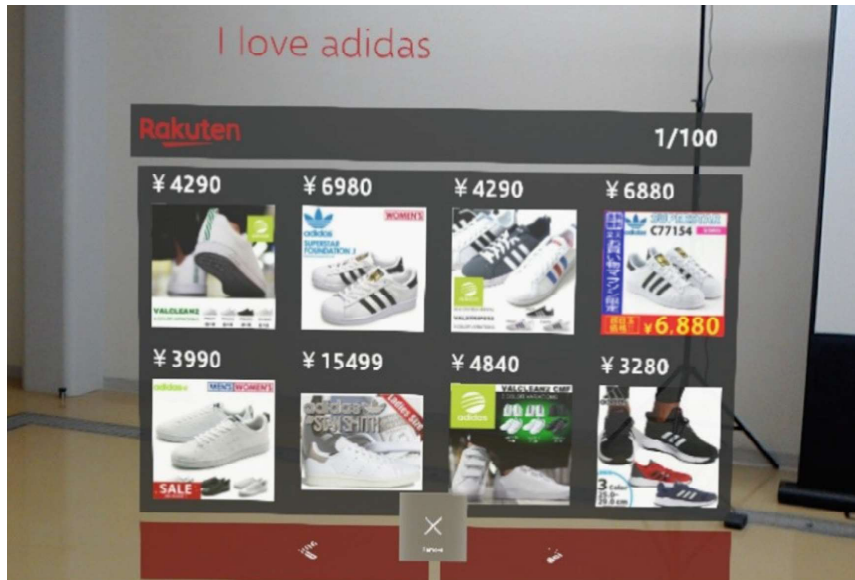


Fig. 3. Users can filter the sneakers by saying "I love adidas."

3.3 Immersive Preview

Our system offers two ways of providing users with immersive shopping experience in augmented reality: automatic placement of virtual products, and manipulation of virtual products with two hands.

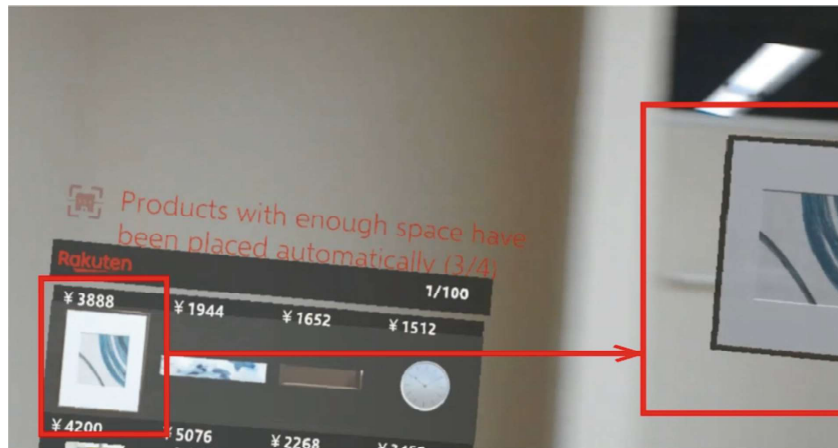


Fig. 4. The system automatically found a place to put the wall painting on the wall

Automatic Placement. In order to let users visually see whether the recommended products are suitable for the current environment, we designed a way to automatically place virtual products into the real world (see Fig. 4).

Before getting the recommended products, we have taken the current scene to extract the dominate objects in the real environment. By recording the location of these key objects, we can place the recommended items in a physical location nearby. Different products have their own placement rules. For example, decorative paintings and racks are usually hung on the wall. The vases are usually placed on a table or other flat surface, and chandeliers are usually hung from the ceiling. In order for the virtual object to automatically follow the placement rules for placement, we needed a higher level of understanding about the user's environment.

We solved this problem by introducing the spatial mapping function. Using this, we can analyze the basic structure of the user's current environment and identify surfaces such as walls, ceilings, and floors. We set up a voice command "place". When users invoke this voice command, the system will tell them to walk around, while collecting the surfaces information of their environment by using HoloLens' depth camera and environment-aware cameras. When enough data is collected, the system will notify users to stop scanning, and the spatial understanding component in the Microsoft Mixed Reality Toolkit [7] will analyze the scanned data. Through this analysis, we can get the orientation (vertical or horizontal), position and size of the corresponding surface in 3D space.

The virtual products are pre-registered with features corresponding to the plane attributes. Products can be pre-registered with one of four attributes: "placed on a wall",

“placed on a floor”, “placed on a ceiling”, or “placed on a platform”. We search the environment in which the users are located to see if there is a corresponding surface. Then decide whether to place the products.

In addition to the surface placement rules, different items also have different layout-related placement habits. We designed another set of placement rule. Products can be pre-registered as “placed in the center of a surface”, “placed on the edge of a surface”, “placed away from other virtual objects that in the same surface”, and “placed in a specific coordinate (Close to a real object)”. With this rule, the products can be automatically placed in the appropriate position of the detected surface.

Two-Hand Manipulation. To provide an augmented-reality preview experience, the products are shown as 3D virtual objects. It helps users to preview the products situated in the real-world setting. We use the two-hand manipulation to make the preview more practical. Two-hand manipulation includes two kinds of operations: drag and rotate. Users can use one hand to drag the virtual objects as if they were grasping an object in the real world. We have also implemented the spatial mapping, so users can drag the virtual object to a position on a physical surface, such as the floor or table. When users want to perform the rotate operation, they can use two hands to manipulate the object.

4 Implementation

Figure 5 shows the overview of our system. We divided this system into two parts: HMD client and server.

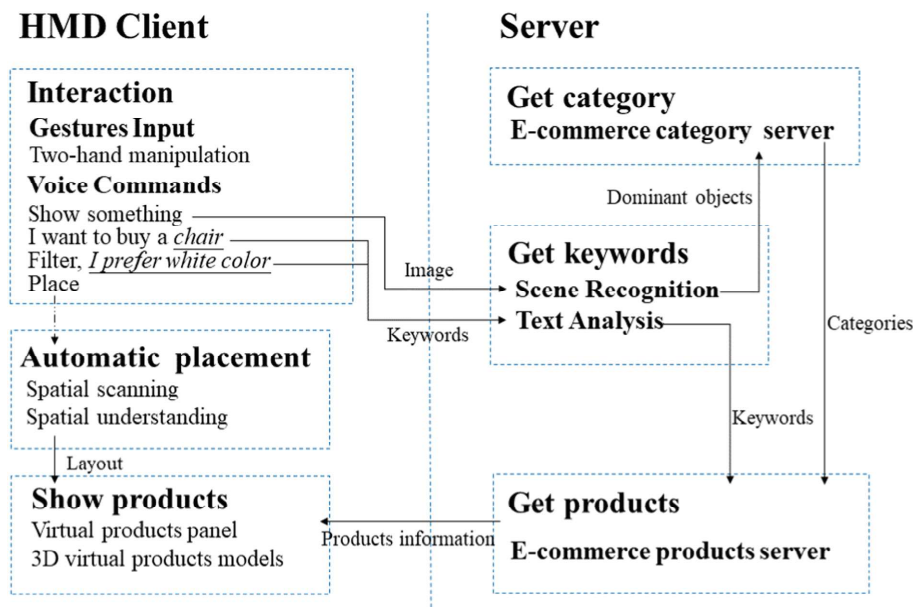


Fig. 5. System overview

On the client side, we used Microsoft HoloLens as the user's terminal. It blends cutting-edge optics and depth sensors to deliver 3D holograms pinned to the real world around the user. Built-in microphones and speakers are used to capture the user's voice and provide audio feedback. When the user makes a voice command, we use the HoloLens dictation system to convert the user's voice commands into text.

4.1 Get Virtual Products

Get Recommended Items Through the Current Scene. When the voice command “show something” is issued, the HoloLens' camera is used to capture the current scene and send the captured photos to the server's Microsoft Azure computer vision server [8]. A description of the current scene and the dominant objects of the scene will be extracted. The current scene-dominant objects are selected, and their related categories are obtained from an e-commerce platform [9]. Then, product search is performed by a fusion of categories and scene description keyword [10] to obtain products recommendations with strong relevance.

Search and Filter Specific Products with Voice Commands. When voice command is used to search a product directly, the microphone will capture the user's voice command, the voice command will first be classified, and is then sent to the cloud analysis platform [11, 12]. The sentence will be analyzed, and keywords extracted. If users issue additional voice command to further refine their filter conditions, additional keywords will be extracted from the command. These keywords will be sent to the e-commerce platform, and products information will be returned. If the user says “filter”, we will send the current keyword and the filter condition together to the e-commerce platform. In this way, users' searching results will be more and more precise.

4.2 Automatically Place Virtual Products

When the voice command “place” is issued, spatial understanding will begin. The system prompts users to walk around and start scanning the surrounding physical environment. Within the Microsoft HoloLens, the cameras group in front of the device are used to perceive the surface information in the surrounding environment. MRTK's spatial understanding component [8] is used to analyze the metadata captured by the cameras group to identify specific ceilings, floors, walls, and other flat information. Combined with the preset information for the virtual item placement feature, we calculate the location that is suitable for placing the virtual products and automatically place them.

4.3 Interact with Virtual Products

We built a 3D virtual object library to store the 3D models. Each product in our system has its own 3D virtual object. So that when the products result is returned, the 3D virtual objects can also be shown.

We pre-processed the virtual models to make sure that they can be operated successfully. Using MRTK [8], bounding box around the virtual objects is used to help users judge the size of the objects. By adding two-hand manipulation scripts in MRTK, users can use their hands to drag and rotate the virtual objects [13].

5 Evaluation

In this section, we introduce our user research and results analysis. We asked participants to accomplish their shopping tasks in a traditional online shopping mobile application which supports searching for items using text and filtering items by price, sales, and shelf date [14] and our intelligent shopping assistant system (they have the same product library). The main purpose of this study was to test whether our system can provide users with high-quality product recommendations that could interest them and whether our system can help them make shopping decisions in a specific situation. We will also discuss the received feedback from a questionnaire.

5.1 Participants

We invited 12 participants (4 females and 8 males), ranging from 19 to 25 years of age. All participants have basic computer skills. Two of them had experience with head-mounted displays. They were divided into two groups evenly.

5.2 Task and Procedure

Before each study, we introduced the basic operations of Microsoft HoloLens to the participants. After the participants became familiar with the device, we asked them to search for products of interest that are related the current environment in three different scenes (kitchen, student office and one corner of the room). For each scene, group 1 was asked to first use the traditional mobile shopping system for 30 min, and then use our intelligent shopping assistant system to shop for another 30 min (see Fig. 6).

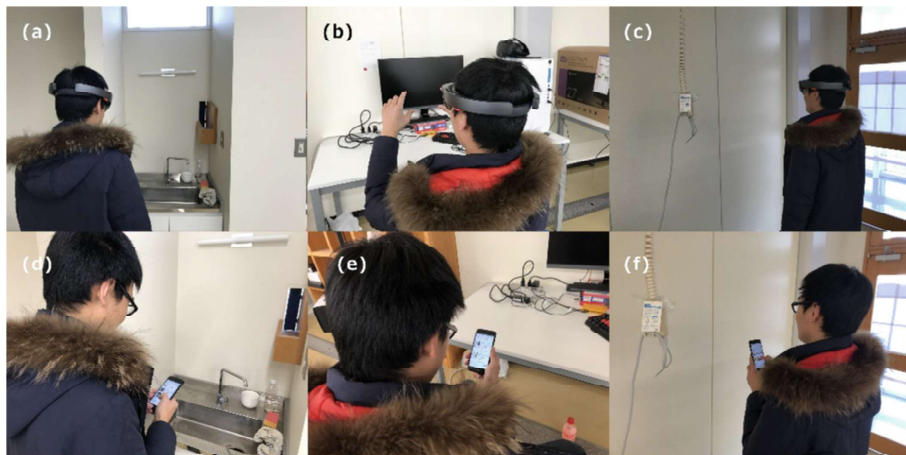


Fig. 6. Evaluation: (a) Using our system in the kitchen. (b) Using our system in the student office. (c) Using our system in one corner of the room. (d) Using current mobile shopping system in the kitchen. (e) Using current mobile shopping system in the student office. (f) Using current mobile shopping system in one corner of the room.

Group 2 was asked the opposite and used our system first. We asked all 12 participants to record the number of products they are interested in when using the shopping system. After the participant completed the whole study, we asked them to fill out a questionnaire with 5 questions to obtain qualitative feedback. Participants rated each question from 1 to 5 (1 = very negative, 5 = very positive).

5.3 Result and Discussion

At the end of the experiment, we collected 36 sets of data from 12 participants using two systems with three scenarios. We calculated the average number of products that participants were interested in (see Table 5).

Table 5. The average number of products that participants interested in thirty minutes.

	Scene	Group1		Group2	
		Current system	Our system	Current system	Our system
1	Kitchen	2.3	4.3	3.0	4.5
2	Student office	1.5	3.3	3.2	5.3
3	One corner of the room	1.0	2.5	2.3	4.8

As shown in Table 5, in the three different scenarios, participants of both groups using our system found more products that were of interest compared to using the traditional mobile e-commerce system. In other words, our system recommends more related items for the current scene for the participants. One participant mentioned that using our shopping system has brought him more shopping inspiration. *“When I use the mobile shopping app in the corner of the room, I don’t have any shopping ideas. When I scanned the environment with your system, I realized that I could buy some paintings to decorate the walls.”*

Table 6 shows the results of our questionnaire. We also divided the results into two parts – using current mobile e-commerce app and using our system. We calculated the average score of each question.

Table 6. Questionnaire.

	Question	Current system	Our system
1	I could get high-quality product recommendation	2.8	4.6
2	I could fully understand the product information	3.8	4.5
3	I felt the system operation was easy enough	4.0	4.0
4	I could easily filter out the products I want	3.0	4.3
5	I could make shopping decisions quickly	2.6	4.7

Question 1 is used to judge whether our system can recommend high-quality products to users. From the result, we can see that participants are not satisfied with the recommendation mechanism of the traditional mobile shopping system under scene-based shopping. This may be because the traditional recommendation mechanism

cannot judge what items are missing in the current scene, so it recommends to users many duplicate items. By using search for products by scene feature of our system, users can get recommended products that are strongly related to the current scene, thereby improving the quality of recommended products.

Question 2 is used to compare traditional mobile shopping application with our system for product information display. As can be seen from the results, participants thought that using our system can better understand products information than traditional mobile shopping systems. Traditional mobile shopping systems typically use limited 2D images and captions to describe the specifics of the product, and users may not be able to get more specific information from these descriptions. Our system allows users to understand the product more fully by visually displaying a rotatable virtual 3D product with descriptive information.

Question 3 is used to test the ease of use and usability of the system. Participants gave us the same rating as the traditional mobile shopping system. The result show that even though users are not familiar with the operation of the head-mounted display, our voice-based system is still easy to use.

Questions 4 and 5 are used to determine whether our system can help users filter out the products they want and help them make purchasing decisions. Our ratings are higher than the traditional mobile shopping system. When using the traditional mobile shopping system, users often need to constantly view the descriptions of the products to make shopping decisions. In our system, by using voice commands, users can easily narrow down the search range and filter out items of interest. With the automatic placement feature, users can check whether a virtual item has a suitable location to be placed and visually compare different virtual items in the environment. This in turn can make shopping decisions faster.

In general, all participants rated our system higher than the traditional system. This may signify that our system is designed to be reasonable and practical. It demonstrates that our system can build a new way of shopping, allowing users to intuitively get high-quality recommended products based on the current scene and make quick shopping decisions.

6 Related Works

One of the related works is a smart assistant toward product-awareness shopping [15]. This research employs sensor techniques in developing a smart assistant for home furniture shopping. In their work, it helps consumers to locate the product easily. Their recommendation method eliminated duplicated products display. The assistant can help to avoid the unnecessary crashing of huge shopping carts in a crowded situation. Their work has integrated the consumer, retailer and the warehouse sides. It is a new shopping pattern.

Another related work is MR-Shoppingu [16]. This research realized an MR shopping system which combines physical and digital information spaces to make digital content accessible directly on physical objects. In this research, they also used Microsoft HoloLens as the see-through type head-mounted display. When the user sees products through

the head-mounted display, he can see the products as well as the digital information of the products which includes pricing information, customer reviews or product video.

Their research provides users a natural way of interacting with the real world without any special input. The combination of the real world and the digital world can provide a more entertaining shopping experience for consumers. This work has inspired our design of virtual user interface to help us build an immersive feeling to the user.

7 Conclusion and Future Work

In this paper, we presented an intelligent shopping assistant which support quick scene understanding and immersive preview. By integrating cloud services, the system can smartly recommend products which are related to the current scene. Users can search for products with natural language using voice command and invoke multiple modification conditions and filter search results. After receiving the results, the system can help users place the virtual objects automatically in their scene and enabling users to manipulate virtual products with hand gestures.

We received feedback from the users by conducting a user study. Overall, we received positive feedback and our system is practical that the users are willing to use it.

In future work, we will continue to improve the system. Since the system is in a demonstration stage, the system is not fully completed. The number of 3D virtual models is very limited. Only some of the objects have 3D virtual objects, so in the future work, it is necessary to expand the library of 3D virtual products.

References

1. Luo, P., Yan, S., Liu, Z.Q., Shen, Z.Y., Yang, S.W., He, Q.: From online behaviors to offline retailing. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2016, pp. 175–184 (2016)
2. Westus.dev.cognitive.microsoft.com: Microsoft Cognitive Services Text Analytics API (v2.0) (2018). <https://westus.dev.cognitive.microsoft.com/docs/services/TextAnalytics.V2.0/operations/56f30ceeda5650db055a3c6>. Accessed 10 Jan. 2019
3. Docs.microsoft.com: Gestures - Mixed Reality (2018). <https://docs.microsoft.com/en-us/windows/mixed-reality/gestures>. Accessed 9 Jan 2019
4. O'Hara, K., Perry, M.: Shopping anytime anywhere, extended abstracts. In: Proceedings of the ACM CHI 2001 Conference on Human Factors in Computing Systems, pp. 345–346 (2001)
5. Li, L.-J., Li, F.-F.: What, where and who? classifying events by scene and object recognition. In: 2007 IEEE 11th International Conference on Computer Vision, pp. 1–8 (2007)
6. Wang, Y.Y., Yu, D., Ju, Y.C., et al.: An introduction to voice search. *IEEE Signal Process. Mag* **25**(3), 28–38 (2008)
7. Docs.microsoft.com: Describing images - Computer Vision - Azure Cognitive Services (2018). <https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/concept-describing-images>. Accessed 13 Jan 2019

8. Microsoft 2018b: MixedRealityToolkit-Unity, 29 March 2018. <https://github.com/Microsoft/MixedRealityToolkit-Unity>. Accessed 12 Jan 2019
9. Webservice.rakuten.co.jp: Rakuten Web Service: Ichiba Genre Search API (version:2012-07-23)|API (2019). <https://webservice.rakuten.co.jp/api/ichibagenresearch>. Accessed 12 Jan 2019
10. Webservice.rakuten.co.jp: Rakuten Web Service: Ichiba Item Search API (version:2017-07-06)|API (2019). <https://webservice.rakuten.co.jp/api/ichibaitemsearch/>. Accessed 12 Jan 2019
11. Docs.microsoft.com: Voice design - Mixed Reality (2018). <https://docs.microsoft.com/en-us/windows/mixed-reality/voice-design>. Accessed 15 Jan 2019
12. Docs.microsoft.com: Voice input - Mixed Reality (2019). <https://docs.microsoft.com/en-us/windows/mixed-reality/voice-input>. Accessed 17 Jan 2019
13. Chaconas, N., Hollerer, T.: An evaluation of bimanual gestures on the microsoft hololens. In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 33–40 (2018)
14. Mobile APP|Rakuten.com (2019). <https://www.rakuten.com/info/mobile-app/>. Accessed 15 Jan 2019
15. Chen, C.C., Huang, T.C., Park, J.J., et al.: A smart assistant toward product-awareness shopping. *Pers. Ubiquit. Comput.* **18**(2), 339–349 (2014)
16. Cheng, K., Nakazawa, M., Masuko, S.: MR-Shoppingu: physical interaction with augmented retail products using continuous context awareness. In: Munekata, N., Kunita, I., Hoshino, J. (eds.) ICEC 2017. LNCS, vol. 10507, pp. 452–455. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66715-7_61